DOCUMENT RESUME

ED 408 361 TM 026 761

AUTHOR Ashby, John
TITLE Standardised Assessment of Mathematical Concepts in Students Aged Four Plus in England and Wales.
PUB DATE Mar 97
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Child Development; *Educational Assessment; Foreign Countries; Formative Evaluation; *Mathematical Concepts; Mathematics Tests; Pilot Projects; *Preschool Children; Preschool Education; School Readiness; *Standardized Tests; Surveys; *Test Construction
IDENTIFIERS *England; *Wales

ABSTRACT

The requirement that schools in England and Wales assess their students, at age seven, in English and mathematics has created a demand for an assessment of students at their entry to statutory schooling. This paper reports on the development and standardization of a test of mathematical concepts for students aged four and over (four plus) as they enter school. The test would give baseline information about the students and would help in formative judgments about teaching. Because students at this age have very limited reading and fine motor skills, it was necessary to develop a test in which all items are spoken by the teacher and anything the student records is within their capabilities. A national mail survey of 122 local education authorities and a telephone survey of 20 schools resulted in a domain range for the first version of the assessment. Trials of 2 versions of the test with a total of 483 children resulted in a version that was tested with a standardization sample of 279 schools in England and Wales in 1994. Results from 1,749 students and the opinions of 72 teachers supported the usefulness of the test, and provides numerical outcomes that can be used for value-added analysis. (Contains 2 figures and 16 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY
John Ashby
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

ED 408 361

STANDARDISED ASSESSMENT OF MATHEMATICAL CONCEPTS IN STUDENTS AGED FOUR PLUS IN ENGLAND AND WALES

A Paper for the AERA Annual Meeting
Chicago 1997

John Ashby
National Foundation for Educational Research
The Mere
Upton Park
Slough
SL1 2DQ
England

BEST COPY AVAILABLE

TM026761

Standardised assessment of mathematical concepts in students aged four plus in England and Wales

Introduction

The statutory requirement for schools in England and Wales to carry out assessments of their students aged seven, in English and mathematics has fuelled a demand for the assessments of students at the age of four plus at the time of their entry to statutory schooling. This paper reports on the development and standardisation of a test of mathematical concepts of students aged four plus.

Background

The Education Reform Act 1988 instigated the introduction in 1991 of national assessments in English and mathematics for all students of seven years of age in England and Wales. The government's stated purpose for doing this was to 'show what a pupil has learnt and mastered, so as to enable teachers and parents to ensure that he or she is making adequate progress and to inform decisions about next steps' (GB. Statutes 1988). Although these data from the assessments provided schools and Local Education Authorities with useful information, a great deal of concern has been expressed about the interpretation of the results particularly when schools and Local Education Authorities are required to disseminate the outcomes of the national assessments as part of a political drive to increase their public accountability. The concern arises because differences in national assessment results do not take social factors into account. Any group of students entering statutory schooling will have had a variety of experiences with mathematical ideas. The extent to which these ideas have developed will depend on the teaching and learning situations they have experienced with their families and peer groups. Some may have been given more structured experiences in nursery schools.

Since 1991 there has been a growing demand in England and Wales for assessments of students' attainment in English and mathematics to be carried out as they enter statutory schooling at the age of four plus. This is so that the outcomes of these assessments can be used as a baseline marker against which students' subsequent performance at seven years of age can be compared. In this way a measure of the value-added by a school can be made.

Currently, a majority of schools in the country carry out some form of assessment of English, mathematics and personal and social skills on their students aged four plus. These assessments usually take the form of teachers' observations of students'

performance in class. These observations are sometimes augmented by parents' observations of a student's performance at home. Teachers usually record the outcome of their observations on a tick list of criteria or as a short discursive report; sometimes they do both. Teachers use this information for making formative judgements when planning interventions or curriculum planning more generally. There is an increasing minority of schools that use a standardised test instrument in conjunction with their observation-based assessment.

A national scheme of baseline assessment is being piloted in September 1997 and from September 1998 all schools in England and Wales will be required by law to carry out assessments in English, mathematics and personal and social development of their students, aged four plus at the time of, or shortly after, their entry to statutory schooling.

Local Education Authorities will be permitted to develop and use their own procedures for baseline assessments provided that such procedures satisfy the criteria for accreditation as set out by the government. Thus, there will be no requirement for schools to use any specific tests or procedures, nevertheless, all procedures will have to be accredited. A major criterion for accreditation is that a baseline assessment procedure must 'provide one or more numerical outcomes, capable of being used for value-added analysis'. (GB, Government Consultation Document 1997).

The purpose of the research set out in this paper was to develop a test of mathematical concepts of students aged four plus which could give teachers detailed information that would be useful when making formative judgements and in addition would give numerical information that could be used for baseline markers. Such an instrument could be used by schools and Local Education Authorities as part of their overall baseline assessment strategy and in conjunction with any existing procedures.

The challenge was to develop an appropriate set of mathematical items, to form a test with acceptable statistical properties and to standardise it on a nationally representative sample of students.

Domain specification

The national assessment tests for students aged seven have their content defined by the domains that are specified in the national curriculum for England and Wales (GB. DFE and WO, 1995). There is no similarly detailed national domain specification for students aged four plus. Therefore the development of test items had to be prefaced

by the construction of a domain specification and sensitive value judgements had to be made about a suitable specification for students of this age.

Initially, there were three sources of information that could be used to make such judgements. These were the findings of research on young students' mathematical skill and understanding, a survey of the mathematical domains that schools in England and Wales were currently using to assess their students aged four plus and an extrapolation of the mathematical domains for students aged five plus that were laid down in the national curriculum. A fourth point of reference then became available, a national specification for the performance outcomes of students leaving nursery schools (GB. SCAA and DFEE 1996).

Research on what young students know about mathematical ideas was once dominated by the work of Piaget (Piaget 1952, Piaget 1953, Piaget & Inhelder 1956, Piaget et al 1960). Whether children conserve number, length, area and so on is still an issue. However, more emphasis is now placed on the skills which young students develop before they enter statutory schooling. These are not always fully recognised upon entry to school and they may come into conflict with the more formal learning they are expected to engage with in their classes (Macnamara, 1992; Aubrey 1993; Aubrey 1994a, 1994b; Munn, 1994). These abilities include: perceiving the number of objects in a small group (Macnamara, op cit); representing number (Hughes, 1986); adding and subtracting (Carpenter and Moser, 1984); dividing/sharing (Davis and Pepper, 1992); geometrical thinking. This work stresses the importance of two main principles: 1) teachers starting where a student is in terms of knowledge and skills, and 2) that pre-school students develop their own strategies for dealing with mathematical problems. These findings provided pointers to both the content and style of item in developing an assessment instrument for students aged four plus.

Further indications for the content of items were obtained from a survey of the domains that were currently being used to assess students aged four plus in schools in England and Wales and also by extrapolating the domains in the specification of the national curriculum for mathematics for students aged five plus. By utilising all of these sources of information the contents specification for the test was devised.

Appropriateness of mode of assessment

Students of this age have a very limited range of reading skills and fine motor skills. It is therefore essential that all instructions and questions are spoken by the teacher and that any recording that the student is required to do is simple enough to be within

the capability of their fine motor skills' development. Furthermore, the content of instructions and questions must be such that young students understand what is being required of them. The task must have meaning to students of this age.

The assessment instrument

The domain coverage and assessment criteria are set out in the accompanying extracts from the Teacher's Guide headed 'Assessment Focuses'. Reference to this extract will indicate that the concepts and skills being assessed are those within the domains of number, pattern, measurement, shape and space. Within these domains there are varying numbers of items. The domain of number has the most items: 38, and measurement has the fewest: 2. In total there are 55 items in the instrument but these do not have to be completed during a single administration. Each item is presented by the teacher as a very short task. Items are grouped together in domain areas or by task type. Each group of items can be completed as a stand-alone administration.

In the course of the range of assessments that focus upon number skills the student is asked to identify number symbols, recite number names in correct order, count pictorial representations of quantities of stars, snails, butterflies, flowers and buttons and count physical objects such as coins and cubes. Students can indicate their responses in any way that is suitable for example by pointing to number symbols or by making wax crayon marks on number symbols, by physically stacking cubes or by touching each cube given.

When students are being assessed on their skills relating to pattern they are asked first to copy a simple repeating pattern of circles and triangles and then to continue a simple repeating pattern of circles and squares. Students can do this by either drawing shapes or by sequencing plastic geometric shapes.

In the assessment of measurement students are presented with a simplified pictorial representation of a worm and asked to draw one that is longer and another that is shorter and when being assessed on their skills in shape and space, students are asked to identify simple geometric shapes.

The Pupil's Booklet is in colour. It is intended for use by both the student and the teacher. In many parts of the Booklet it is the teacher and not the student who is recording the student's responses. In other parts it can be either the teacher or the student. There are only two items in the whole instrument where it has to be the student doing the recording and these are the two measurement items. There is

provision on each page of the Booklet for teachers to make notes about students' strategies and misunderstandings. Thus, the Pupil's Booklet has a wider use than as a test booklet; its role is a combination of test and recording instrument.

At the conclusion of working through the assessments, teachers will have standardised scores and detailed information on the domain-specific achievements of the students within the domain coverage of the instrument. In addition to this they will have any notes that they have made on students' strategies and misunderstandings. Extracts from the Pupil's Booklet are attached.

The development

In order to obtain information about the point-of-entry assessment practices of schools with particular reference to domain coverage of early mathematical concepts, a national postal survey was carried out involving all 122 Local Education Authorities in England and Wales and a telephone survey of a nationally representative sample of 20 schools. The outcome of this survey indicated a domain range that was incorporated into our domain specification.

Initially, individual items were developed by researchers who carried out limited informal trials in schools with students at the appropriate age. Items were subsequently combined to make two versions of the assessment instrument. Both versions were trialled with a nationally representative sample involving 483 students in total. Two versions of the instrument were trialled in order to facilitate the creation of a large pool of trialled-items. Both instruments assessed the same domains but contained items that varied in the type of stimulus material given, in the graphical presentation of similar stimulus material or in the actual examples presented. A few items were trialled as parallel versions.

All of the teachers in these trials were asked to administer the assessments to eight students who were in their first six weeks in school. The majority of the items were administered on an individual basis but there was a sample of teachers who reported that they had found it possible to administer some of the items with groups of two or three confident students.

Version A of the test instrument was administered to 255 students. It had 69 single-mark items. The mean score from the trials was 47.3, the standard deviation (SD) 14.3 and the range was 10 to 69. The test's internal consistency was measured by Cronbach's alpha and was 0.96.

There were 228 students in the sample who took test Version B. This Version had 71 single-mark items, a mean score of 52, a standard deviation (SD) of 13.6 and a range of 7 to 71. Cronbach's alpha for this test was 0.95.

The standardisation

The standardisation exercise was carried out during the first three weeks in October 1994 in a nationally representative sample of 279 schools in England and Wales. The sample was proportionately stratified to give a more accurate representation of the country as a whole.

In most of the schools it was the class teacher who carried out the assessments but in a few schools the test was administered by the head teacher or by a support teacher.

The test had to be completed within the trialling period of three weeks and because of the restricted time scale each teacher was asked to administer the test to a total of seven students. In order to achieve an unbiased sample, the criterion by which teachers made the choice of these seven students was specified.

A total of 1749 students were assessed. There were 897 boys and 848 girls and 4 students whose gender was not recorded. The students' ages ranged from 3 years 10 months to 5 years 8 months. The average age was 4 years 8 months and the majority of pupils' ages ranged from 4 years 2 months to 5 years 1 month.

The test was constructed from the pool of items that had been trialled. The instrument contained 55 items in total: 50 were single-mark items; there were two that were double-mark items, two that carried a maximum of four marks and one item that carried a maximum of five marks. Thus, the total possible score was 67. The mean raw score for the standardisation sample was 42.4 and the standard deviation (SD) was 13.5. The range of scores was one to 67. The internal consistency of the test as measured by Cronbach's alpha was 0.91.

Age differences in the national standardisation sample were computed for students in two-monthly age bands and were as follows.

Age range (years and months)	Number of students	Mean	SD
younger than 4.2	13	40.8	11.1
4.2-4.4	217	33.9	13.9
4.5-4.7	282	40.5	12.9
4.8-4.10	494	41.8	12.9
4.11-5.1	695	45.8	12.9
older than 5.1	20	51.2	10.5
age unknown	28	45.5	13.0

In accordance with usual British practice, standardised scores were computed with a mean of 100 and standard deviation of 15 and an adjustment for age was included. The results were also grouped into quintiles.

Item statistics showed a range of facilities from 92 right down to single figures. At the top end of the range, most students in the sample found the following items easy:

- reciting numbers from one to five
- identifying which of two sets contained the most where both sets contained less than ten
- identifying a circle, a square, a triangle
- drawing a line that was shorter than one given

and at the bottom end of the range most students in the sample found the following items difficult:

- reciting numbers from one to 20/30/40
- copying patterns with more than two shapes (allowing for motor skill limitation)
- continuing patterns with more than two shapes (allowing for motor skill limitation).

In addition to the Pupil's Booklets and the Teacher's Guide, each teacher in the standardisation sample received a four page questionnaire requesting details about the management and time taken for the test's administration. Teachers were also asked in the questionnaire to comment on various aspects of the test.

Almost all of the teachers administered the test to students individually. Seven out of ten teachers took an average total time of 30 minutes or less to complete the assessments for an individual student. This total average time was the result of 1) a single administration of the whole test to an individual student or 2) repeated administrations of different parts of the test to an individual student or 3) the combination of individual and group administrations of the test to an individual student.

There were 72 teachers who administered the test to a small group of students on some occasions. Some teachers commented that initially they attempted to work in a small group but quickly abandoned this in favour of working with individual students. There were two reasons why teachers felt the need to work with students individually. Firstly, the restricted language skills of students of this age meant that teachers felt the need to give instructions on a one-to-one basis in order to be confident that the student understood what was being required of him or her. Secondly, students' responses were biased by those of other students in the group. The inquisitiveness of young students meant that they tended to be interested in the responses of other students and to change their own responses if they were different from those of other students. However, teachers commented that they liked working with students individually and the students enjoyed the attention.

Where group administration worked it was usually with two confident and more able students. There were a few teachers who reported that they administered most of the test to groups of four or more students with no undue problems.

As the majority of tests were administered to the students individually the time taken to assess each student was a concern of many teachers. There were two aspects to this concern: 1) If teachers needed to give concentrated attention to individual students whilst at the same time maintaining effective management of the rest of the class then the time taken for each individual administration needed to be short; 2) If all of the students in a class were going to be tested on an individual basis then the total administration time could be prohibitive.

The majority of teachers indicated that the information that they obtained from the test was both illuminating and informative.

Concluding remarks

Teachers should find the test useful because it could form part of an accredited scheme that will meet what will be a statutory requirement in England and Wales from September 1998 onwards. It will provide numerical outcomes that can be used for value-added analysis. It will also give detailed information about students' performances within domains. Teachers should find it helpful to have the flexibility in how they can administer the test: for example whether they work in small groups or with students individually, whether they administer all of the items at one go or break the test up into several administrations. However, issues remain about the administration of any such assessments with students of this age.

The needs of young students dictate that most teachers will want to carry out administrations with students individually and this has implications for the practicability of assessing a whole class. The current practice in most schools in England and Wales is for the teacher to make observations of students' performances in class. This tends to be carried out during a period of several weeks and when assessments of these students becomes a statutory requirement schools will still be allowed several weeks in which to complete the assessment. The problem with this is that if a student is assessed well into the school term, the performance that the student exhibits is not the level of knowledge and understanding that he or she has brought to school as the outcome of prior experience but increasingly is the expertise that has been gained from being in the school. In other words it is no longer a point-of-entry assessment. We recommend that our test is used during a student's first six weeks in school. However the issue comes back to one of practicability of administration for a whole class. In our experience of the various triallings of our test instruments, teachers are prepared to accept a considerable time commitment to the assessments and a reorganisation of their timetables if they perceive that the process is worthwhile in terms of experience for the student and in the type and quality of the information that the assessments provide. Nevertheless, these teachers were carrying out assessment on a voluntary basis. It may be that when it becomes a statutory requirement they will feel that the imposition is such that extra resourcing, in the form of a temporary extra teacher, is necessary. This was the phenomenon that occurred with the national testing for students aged seven.

A different issue is whether it is reasonable to use the results of two different tests for value-added analysis. In order to do so, there should be a clear and definite relationship between the two instruments. Ideally both instruments should have the same domain coverage but this is a problem because the content of the domain for

students aged four plus is inappropriate for students aged seven and vice versa. Further research is needed in this area.

Other questions remain about the efficacy of carrying out standardised assessments at all on such young students. The interpretation of what is typical performance has to be approached cautiously because young students tend to give idiosyncratic responses. This has implications for a test's reliability. However, it was interesting that the various versions of our instrument achieved values for Cronbach's alpha greater than 0.9 as a measure of internal consistency. This was higher than we had anticipated at the outset.

Other concerns about the efficacy of such assessments hinge upon the potential for the inappropriate use of the test results. Such concerns impact upon who should have access to such information, the level of detail provided in the information and the understanding that teachers and parents have about the interpretation of the results. Further research is needed on how different teachers of young students in schools in England and Wales use assessment information.

Publication

The test is due to be published by NFER-NELSON April 1997.

Enquiries to:

NFER-NELSON
Darville House
2 Oxford Road East
Berkshire
Sl4 1DF
England
Tel: +44 (0)1753 858961
Fax: +44 (0)1753 856830

References

GREAT BRITAIN. STATUTES (1988). *Education Reform Act 1988. Chapter 40.* London: HMSO.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION and WELSH OFFICE (1995). *Mathematics in the National Curriculum.* London: HMSO.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY and GREAT BRITAIN. DEPARTMENT FOR EDUCATION AND EMPLOYMENT (1996). *Nursery Education, Desirable Outcomes for Children's Learning on Entering*

Compulsory Education. London: HMSO.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (1997). *Draft Criteria and Procedures for Accreditation of Baseline Assessment Schemes.* London: HMSO.

AUBREY, C (1993) 'An investigation of the mathematical knowledge and competencies which young children bring into school', *British Educational Research Journal,* **19**, 27-41.

AUBREY, C (1994a). 'An investigation of children's knowledge of mathematics at school entry and the knowledge their teachers hold about teaching and learning mathematics, about young learners and mathematical subject knowledge', *British Educational Research Journal,* **20**, 1, 105-120.

AUBREY, C (1994b). 'The construction of early years mathematics'. In: AUBREY, C. (Ed) *The Role of Subject Knowledge in the Early Years of Schooling.* London: Falmer Press.

CARPENTER, T. and MOSER, J. (1984) 'The acquisition of addition and subtraction concepts in grades one through three', *Journal for Research in Mathematics Education,* **15**, 3, 179-202.

DAVIS, G. and PEPPER, K. (1992) 'Mathematical problem solving by pre-school children', *Educational Studies in Mathematics,* 23, 4, 397-415.

HUGHES, M. (1986) *Children and Number: Difficulties in Learning Mathematics,* Oxford: Blackwell.

MACNAMARA, A (1992) 'How many are there?' *Early Years,* **13**, 1, 21-23.

MUNN, P. (1994) 'The early development of literacy and numeracy skills', *European Early Childhood Education Research Journal,* **2**, 1, 5-16.

PIAGET, J (1952) *The Child's Conception of Number.* New York: Humanities Press

PIAGET, J (1953) 'How children form mathematical concepts', *Scientific American,* **189**, 5, 74-79.

PIAGET, J. and INHELDER, B. (1956) *The Child's Conception of Space.* London: Routledge & Kegan Paul.

PIAGET, J., INHELDER, B. and SZEMINSKA, A. (1960) *The Child's Conception of Geometry.* New York: Basic Books.

Extracts from the Teacher's Guide

ASSESSMENT FOCUSES

Number

Skill	Assessment Criteria	Question Numbers
■ Number Terms		
Associate sound with symbol	*Identifies numerals when told name.*	24-27
	States name when numeral is presented.	9-12
Recall names in order	*Recites number names starting at 'one' and in the correct order, going as high as he or she is able to.*	1
▲ Counting		
Associate sound with quantity	*When told name, matches to set of appropriate size.*	37-38
	When presented with a set, states how many the set contains.	7-8 4-5
Associate symbol with quantity	*When presented with numeral, matches to set of appropriate size.*	34-36
	When presented with set, matches to numeral.	34-36
Counting quantity less than 100	*Counts a given set.*	19, 21 4-5
Counting on	*Having counted set, continues count when further items added.*	20
	Names preceding or succeeding number to number given.	2-3
Identify quantities	*Having counted two sets, identifies which contains a given number of items.*	40, 42, 44

11

Number

Skill	Assessment Criteria	Question Numbers
● Combining sets		
Counting more than one set	*Given two sets, states total for both (total less than 11).*	6
● Partitioning sets		
Counting and identifying quantities	*Makes two sets of specified size from larger set.*	48-49
	States new total when a number of items are removed from a set of known size.	22-23
● Comparison		
Comparing sizes of sets	*Identifies which of two sets contains more or fewer.*	39, 41, 43
	Identifies which of a series of sets contains the most.	45-47
Comparing numerals	*Identifies which of two numerals is greater.*	28-31
	Identifies which of a series of numerals is greatest.	32-33
Finding differences	*Finds the difference in size between two sets.*	40, 42, 44

12

Pattern

Skill	Assessment Criteria	Question Numbers
❖ Identify and reproduce attributes	*Copies an alternating pattern.*	50
	Copies a more complex pattern.	52
	Continues an alternating pattern.	51
	Continues a more complex pattern.	53

Measurement

Skill	Assessment Criteria	Question Numbers
✱ Single comparison	Identifies which of two objects is longer or shorter	54-55

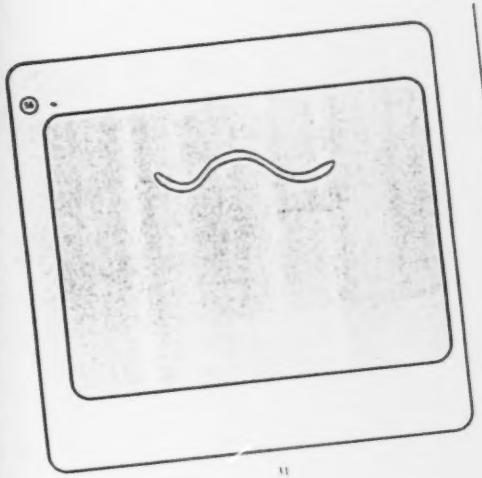
Shape and Space

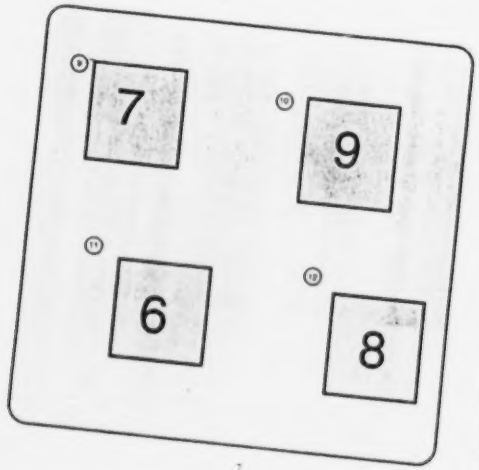
Skill	Assessment Criteria	Question Numbers
★ Associate name with two dimensional shape	*Names square/circle/triangle when presented with shape.*	13-15
	Identifies triangle/circle/square when told name.	16-18

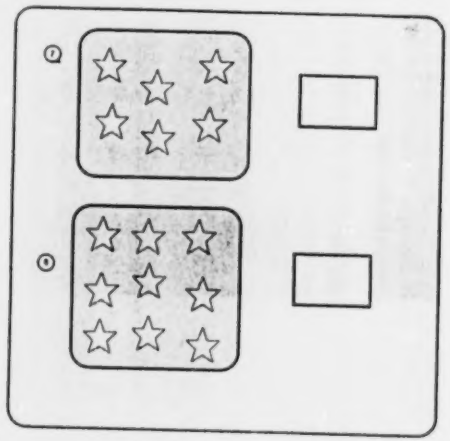
13

BEST COPY AVAILABLE

Extracts from the Pupil's Booklet







BEST COPY AVAILABLE